Aufgabe 1

Für die Securance-Versicherung liegen Ihnen die gemeinsamen absoluten Häufigkeiten der Merkmale X: "Schadenshöhe" und Y: "Versicherungsart" für die letzten 500 gemeldeten Schäden vor.

$y_i \downarrow x_j \rightarrow$	gering	mittel	hoch	$n_{i\cdot}$
Hausrat	200	25	0	225
KFZ	0	10	75	85
Unfall	20	35	90	145
Reiserücktritt	40	5	0	45
$n_{\cdot j}$	260	75	165	500

1. Interpretieren Sie n_{13}.

2. Geben Sie jeweils den Merkmalstyp und das Skalenniveau für das Merkmal X und für das Merkmal Y an.

3. Berechnen Sie für das Merkmal Y ein geeignetes Lagemaß.

Die Securance-Versicherung hat im Durchschnitt 1 Schadensfall pro Tag. Betrachtet wird die Zufallsvariable S_k: "Anzahl der Schäden in k Tagen".

4. Nennen Sie die Annahmen des Poisson-Prozesses.

5. Gehen Sie im Folgenden davon aus, S_k sei Poisson-verteilt mit Parameter λ_k.

   (a) Bestimmen Sie den Parameter λ_5 für die Zufallsvariable S_5 (**Hinweis:** Wenn Sie diese Aufgabe nicht lösen können, verwenden Sie im Folgenden $S_5 \sim \text{Poi}(10)$).

   (b) Mit welcher Wahrscheinlichkeit geschehen innerhalb von 5 Tagen mehr als 5 Schadensfälle?

   (c) Mit welcher Wahrscheinlichkeit geschehen innerhalb von 5 Tagen mehr als 3 aber weniger als 8 Schadensfälle?

(d) Bestimmen Sie Mittelwert und Median von S_5.

6. Gegeben sei in `R` ein Vektor `x`, der eine Stichprobe enthält, und eine positive Zahl `lambda`.

   (a) Welche statistische Größe berechnet der folgende `R`-Code?

   ```
   sum(log(dpois(x,lambda)))
   ```

   (b) Geben Sie einen Befehl an, mit dem Sie in `R` 500 Zufallszahlen aus einer Poisson-Verteilung mit $\lambda = 5$ erzeugen können.

Schmierpapier:

Aufgabe 2

Die Securance-Versicherung modelliert die Zeitdauer zwischen zwei Unfällen mithilfe der Zufallsvariable X: "Wartezeit bis zum nächsten Unfall in Stunden". X sei exponentialverteilt mit Dichtefunktion:

$$f_X(x) = \lambda \exp(-\lambda x), \quad x > 0, \lambda > 0.$$

1. Wie groß ist die Wahrscheinlichkeit, dass die Wartezeit zwischen zwei Unfällen mehr als drei Stunden beträgt, wenn der Parameter $\lambda = 1/2$ ist?

2. Es wird der Maximum-Likelihood Schätzer für den unbekannten Parameter λ gesucht.

   (a) Zeigen Sie, dass die logarithmierte Likelihoodfunktion für den unbekannten Parameter λ gegeben ist durch:

$$\ln L(\lambda; x_i) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

(b) Zeigen Sie, dass der Maximum-Likelihood Schätzer für den unbekannten Parameter λ lautet:

$$\hat{\lambda}_{ML} = \frac{1}{\bar{x}},$$

wobei $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. (Keine Berechnung der zweiten Ableitung nötig.)

3. Sie notieren die Wartezeit zwischen 20 Unfällen und fassen die Daten als Realisationen einer Stichprobe $X_1, \dots, X_{20}$ aus einer exponentialverteilten Grundgesamtheit auf. Anschließend berechnen Sie $\sum_{i=1}^{20} x_i = 40$.

   (a) Tragen Sie die Werte der logarithmierten Likelihoodfunktion für $\lambda = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$ in untenstehende Tabelle ein und zeichnen Sie die entsprechenden Werte in ein geeignetes Koordinatensystem.

λ	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1

(b) Berechnen Sie $\hat{\lambda}_{ML}$ aufgrund der Ergebnisse der Stichprobe.

(c) Der Vorstand des Unternehmens vermutet, dass die Wartezeit zwischen zwei Unfällen mehr als zwei Stunden beträgt. Sie formulieren folgendes Hypothesenpaar:

$$H_0 : \lambda \leq \lambda_0 = 1/2 \quad H_1 : \lambda > \lambda_0 = 1/2.$$

Zur Überprüfung obiger Hypothesen ziehen Sie folgende Prüfgröße heran $T_n = 2\lambda_0 \sum_{i=1}^{n} X_i$. Testen Sie auf Basis des obigen Stichprobenbefundes bei einer Irrtumswahrscheinlichkeit von 5% die Vermutung des Vorstandes. Sie wissen, dass die Prüfgröße T_n unter Gültigkeit der Nullhypothese χ^2-verteilt mit $2n$ Freiheitsgraden ist. Vervollständigen Sie das folgende Testverfahren.

Kritischer Bereich:

Kritische Schranke:

Berechnung der Teststatistik:

Testentscheidung:

4. Geben Sie den Befehl an, mit dem Sie in `R` die Wahrscheinlichkeit aus Teilaufgabe 1. berechnen können.

Gegeben ist Ihnen folgender `R`-Code.

```
x=seq(0,4,,100)
y=pexp(x,0.5)
plot(x,y, type="l")
```

5. Vervollständigen Sie folgende Aussage:
   Mithilfe der obigen Befehle wird die ______________________________ einer Exponentialverteilung mit Parameter ________ gezeichnet.

Aufgabe 3

Bei einem Elfmeterschießen, in dem ein Spieler auf einen Torwart schießt, gibt es für den Schützen nur die Möglichkeiten eines Schusses nach links, rechts oder in die Mitte. Der Torwart kann sich ebenfalls links, rechts oder mittig positionieren. Ein Elfmeter gilt als gehalten, wenn sich der Torwart in Schussrichtung positioniert.

1. Für den Verlauf des Elfmeterschießens gelte folgende Wahrscheinlichkeitstabelle:

			Schussrichtung S des Schützen			
			Links	Mitte	Rechts	
			$S=1$	$S=2$	$S=3$	
Position T des Torwarts	Links	$T=1$				1/3
	Mitte	$T=2$				1/3
	Rechts	$T=3$	1/18	1/18	2/9	1/3
			1/3	1/3		1

   (a) Vervollständigen Sie die Wahrscheinlichkeitstabelle. Zusätzlich wissen Sie, dass
   - der Schütze mit einer Wahrscheinlichkeit von 3/10 in die Mitte schießt, wenn sich der Torwart links positioniert,
   - $P(T=2,\ S=1) = P(T=2,\ S=3)$ gilt.

(b) Berechnen Sie $P(T = 3|S = 1)$. Überprüfen Sie, ob die Ereignisse "Der Schütze schießt nach links" und "Der Torwart positioniert sich rechts" stochastisch unabhängig sind.

Hinweis: Wenn Sie Teilaufgabe 1 (a) nicht lösen konnten, verwenden Sie im Folgenden

$$P(T = 1,\ S = 1) = P(T = 1,\ S = 3) = P(T = 2,\ S = 2) = P(T = 2,\ S = 3) = \frac{1}{18},$$
$$P(T = 1,\ S = 2) = P(T = 2,\ S = 1) = \frac{2}{9}.$$

(c) Berechnen Sie die Wahrscheinlichkeit dafür, dass ein Tor fällt (also dass Schussrichtung des Schützen und Position des Torwarts nicht übereinstimmen).

2. Ein Schütze, der mit Wahrscheinlichkeit 40% trifft, hat nun 5 Versuche.

(a) Welche Annahme muss getroffen werden, damit die Zufallsvariable X_5: "Anzahl der Treffer bei 5 Schüssen" binomialverteilt ist?

Gehen Sie von nun an davon aus, dass die benötigte Annahme erfüllt ist.

(b) Geben Sie die Verteilungsparameter der Zufallsvariable X_5 an.

(**Hinweis:** Wenn Sie diese Aufgabe nicht lösen können, verwenden Sie im Folgenden $X_5 \sim \mathrm{Bin}(n = 9,\ p = 0.25)$).

(c) Wie viele Treffer kann der Schütze erwarten?

Wie hoch ist die Wahrscheinlichkeit, dass der Schütze bei 5 Schüssen

(d) genau 3 Treffer erzielt?

(e) mindestens 4 Treffer erzielt?

3. Ihnen steht folgender R-Output zur Verfügung:

```
> dbinom(x=2, size=5, prob=0.68)
[1] 0.1515192
> pbinom(q=2, size=5, prob=0.68)
[1] 0.1905263
```

(a) Ermitteln Sie anhand des Outputs die Wahrscheinlichkeit dafür, dass ein geübter Elfmeterschütze mit Trefferwahrscheinlichkeit 68% bei 5 Schüssen höchstens 2 Treffer erzielt.

(b) Geben Sie den R-Code an, mit dem Sie die Wahrscheinlichkeitsfunktion der Binomialverteilung $f_{\text{Bin}}(x;\ n = 5,\ p = 0.81)$ zeichnen können.

Schmierpapier:

Aufgabe 4

Für eine Untersuchung von Airlines wurde eine einfache Stichprobe gezogen. Folgende Tabelle gibt X: "Kosten pro Sitzkilometer" und Y: "Gebühr pro Stück Übergepäck" von 8 zufällig ausgewählten Airlines an.

Airline i	1	2	3	4	5	6	7	8
X (in Cent)	4.5	14.9	5.0	13.0	7.4	9.4	7.7	8.5
Y (in Euro)	64	80	83	80	80	96	69	120

1. Berechnen Sie die Stichprobenvarianz für Y, wenn $\overline{y}_8 = 84$ ist. (**Hinweis:** Wenn Sie diese Aufgabe nicht lösen können, verwenden Sie im Folgenden $s^2_{Y;8} = 250$.)

2. Man vermutet, dass die Gebühr pro Stück Übergepäck mit den Kosten pro Sitzkilometer positiv korrelliert. Prüfen Sie die Vermutung, indem Sie den Korrelationskoeffizient der Stichprobe berechnen. Ihnen sei bekannt, dass $s^2_{X;8} = 13.06$ und $\sum_{i=1}^{8}(x_i - \overline{x}_8)(y_i - \overline{y}_8) = 67.1$ sind.

3. Folgende Tabelle gibt die Kosten pro Sitzkilometer von 6 Airlines an, die in `R` als Objekt `x` gespeichert sind. Seien die Kosten X in der Grundgesamtheit normalverteilt. (Stand: 2010)

Airline	SAS	Lufthansa	Air France	Iberia	British Airways	Turkish Airlines
X (in Cent)	13.20	12.80	8.90	7.80	6.50	5.70

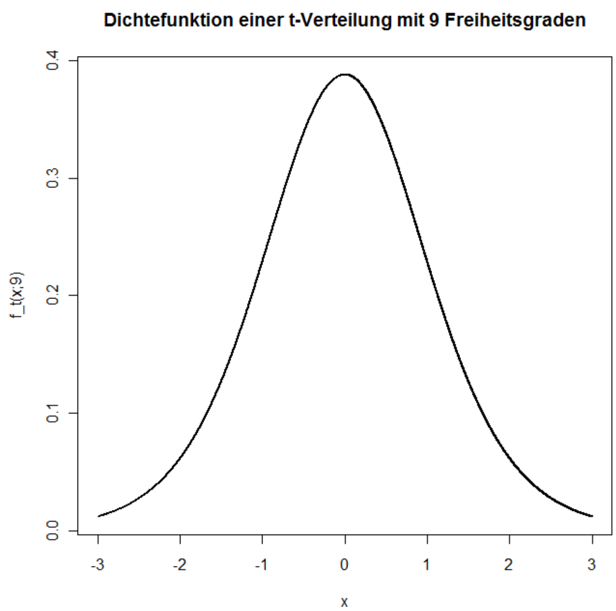
Vervollständigen Sie folgende Aussagen:

(a) Der `R`-Code `qt(0.1,length(x)-1)` berechnet das 10%–__________________ der t-Verteilung mit ________ Freiheitsgraden.

(b)
```
f1=function(x)
{
l=mean(x)-qt(0.95,length(x)-1)*sqrt(var(x)/length(x))
u=mean(x)+qt(0.95,length(x)-1)*sqrt(var(x)/length(x))
c(l,u)
}
```
Die Funktion `f1()` berechnet ein zweiseitiges ________%-Konfidenzintervall für ____ von X.

4. Ein Flugzeug kann maximal 150 Passagiere aufnehmen. Seien $K = 150X$ die Gesamtkosten des Flugzeugs pro km und $X \sim \mathcal{N}(\mu, \sigma^2 = 2)$. Wie und mit welchen Parametern ist K verteilt?

5. Basierend auf einer Stichprobe von 10 Flügen soll auf einem Signifikanzniveau von 10% getestet werden, ob der Mittelwert der Kosten pro Sitzkilometer 12 Cent überschreitet, d.h. $H_0 : \mu \leq \mu_0 = 12$ und $H_1 : \mu > \mu_0 = 12$. Markieren Sie auf folgendem Plot den kritischen Bereich für den dazu passenden Hypothesentest.



6. Auf der Strecke München-Tokyo wird ein neuer Flugzeugtyp eingesetzt, um die Kosten pro Sitzkilometer zu senken. Die Wirksamkeit dieser Maßnahme soll getestet werden. Dazu wurden die durchschnittlichen Kosten von 10 Flügen mit altem Flugzeug und von 10 Flügen mit neuem Flugzeug ermittelt. Folgende Größen liegen Ihnen vor: $\overline{x}_{alt,10} = 10$, $\overline{x}_{neu,10} = 9.5$, gepoolte Stichprobenvarianz $\overline{S}^2 = 15$ und $\sigma^2_{alt} = \sigma^2_{neu}$. Vervollständigen Sie das Testverfahren.

   Hypothesenpaar: $H_0 : \mu_{neu} - \mu_{alt} \geq 0$ gegen $H_1 : \mu_{neu} - \mu_{alt} < 0$,
   Signifikanzniveau: $\alpha = 5\%$.

   Kritischer Bereich:

   Kritische Schranke:

Teststatistik:

Testentscheidung:

7. Vervollständigen Sie folgende Aussagen:
   Die Güte eines Tests wird umso größer,

   - je weiter Null- und Alternativhypothese entfernt liegen;
   - je ____________ der Stichprobenumfang ist;
   - je ____________ die Irrtumswahrscheinlichkeit α ist.

Aufgabe 1

1. $n_{13} = 0$: keiner der Schadensfälle wurde als hoch eingestuft und betrifft eine Hausratsversicherung.
2.
   - Merkmal X, Merkmalstyp: komperativ, Skalenniveau: Ordinalskala
   - Merkmal Y, Merkmalstyp: qualitativ, Skalenniveau: Nominalskala
3. Der Modus liegt bei der Merkmalsausprägung 'Hausrat'.
4. Annahmen des Poisson-Prozesses:
   - Die Wahrscheinlichkeit für einen Ereigniseintritt in einem Intervall hängt von dessen Länge ab, nicht von dessen Lage.
   - Die Anzahl der Ereignisse in zwei sich nicht überlappenden Intervallen sind voneinander unabhängig.
   - Die Wahrscheinlichkeit für mehr als ein Ereignis in einem sehr kurzen Intervall ist vernachlässigbar gering.
5. (a) $S_1 \sim Pois(\lambda_1 = 1) \Rightarrow S_5 \sim Pois(\lambda_5 = 5)$

   (b) $P(S_5 > 5) = 1 - P(S_5 \leq 5) = 1 - F_{Pois}(5;\ 5) = 1 - 0.6160 = 0.3840$

   (c) $P(3 < S_5 < 8) = F_{Pois}(7;\ 5) - F_{Pois}(3;\ 5) = 0.8666 - 0.2650 = 0.6016$

   (d) Erwartungswert: $E[S_5] = \lambda_5 = 5$,
   Median: $F_{Pois}(4;\ 5) = 0.4405 < 0.5$ und $F_{Pois}(5;\ 5) = 0.6160 \geq 0.5$, so dass der Median bei $s_{(0.5)} = 5$ liegt.
6. (a) `sum(log(dexp(x,lambda)))` berechnet den Wert der logarithmierten Likelihoodfunktion für eine Exponentialverteilung an der Stelle `lambda` bei gegebener Stichprobenrealisation `x`.

   (b) `rpois(500,lambda=5)`

Aufgabe 2

1. $P(X > 3) = 1 - F_{Exp(1/2)}(3) = 0.22.$

2. (a)

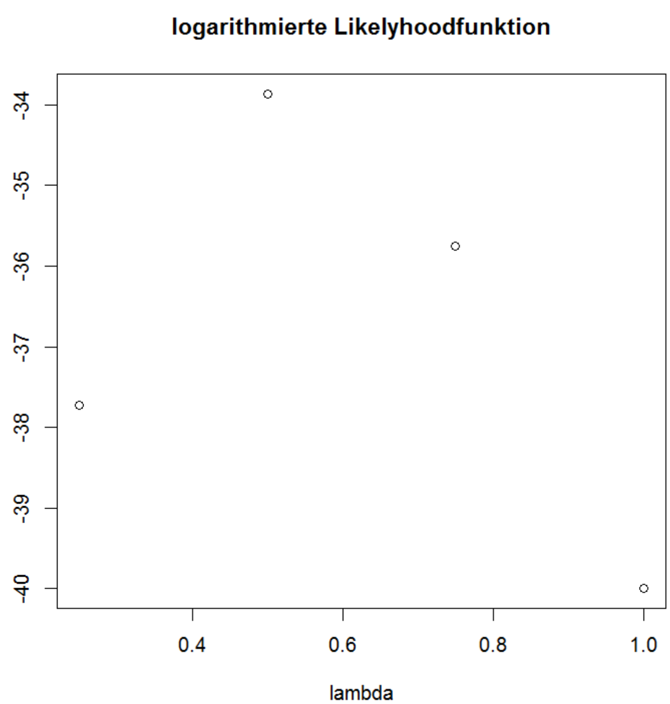
$$\begin{aligned}
L(\lambda, x_i) &= \prod_{i=1}^{n} f(\lambda, x_i) = \prod_{i=1}^{n} \lambda \exp(-\lambda x_i) \\
&= \lambda^n \exp(-\lambda \sum_{i=1}^{n} x_i). \\
\ln L(\lambda, x_i) &= n \ln \lambda - \lambda \sum_{i=1}^{n} x_i.
\end{aligned}$$

   (b)

$$\begin{aligned}
\frac{\partial \ln L(\lambda, x_i)}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^{n} x_i \overset{!}{=} 0. \\
&\Leftrightarrow \frac{n}{\lambda} = \sum_{i=1}^{n} x_i \\
&\Leftrightarrow \hat{\lambda}_{ML} = 1/\bar{x}.
\end{aligned}$$

3. (a)

λ	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1
$\ln L(\lambda; xi)$	-37.72589	-33.86294	-35.75364	-40.00000



(b) $\hat{\lambda}_{ML} = 1/(1/20 \cdot 40) = 1/2$.

(c)
- Kritischer Bereich:
  Falls $T_n > KS$ ist H_0 abzulehnen.
- Kritische Schranke:
  95%-Quantil der $\chi^2(2 \cdot 20)$-Verteilung aus Tabelle ablesen: 55.76=KS.
- Berechnung der Teststatistik:
  $T_n = 2 \cdot 1/2 \cdot 40 = 40$.
- Testentscheidung: $T_n \not> KS$, d.h. H_0 nicht ablehnen.

4. `1-pexp(3,rate=0.5)`

5. Vervollständigen Sie folgende Aussage:
   Mithilfe der obigen Befehle wird die <u>Verteilungsfunktion</u> einer Exponentialverteilung mit Parameter $\underline{\lambda = 0.5}$ gezeichnet.

Aufgabe 3

			Schussrichtung S			
			Links	Mitte	Rechts	
			$S=1$	$S=2$	$S=3$	
Position T	Links	$T=1$	$1/5$	$1/10$	$1/30$	$1/3$
	Mitte	$T=2$	$7/90$	$8/45$	$7/90$	$1/3$
	Rechts	$T=3$	$1/18$	$1/18$	$2/9$	$1/3$
			$1/3$	$1/3$	$1/3$	1

oder

			Schussrichtung S			
			Links	Mitte	Rechts	
			$S=1$	$S=2$	$S=3$	
Position T	Links	$T=1$	0.2	0.1	$0.0\bar{3}$	$1/3$
	Mitte	$T=2$	$0.0\bar{7}$	$0.1\bar{7}$	$0.0\bar{7}$	$1/3$
	Rechts	$T=3$	$1/18$	$1/18$	$2/9$	$1/3$
			$1/3$	$1/3$	$0.\bar{3}$	1

1. (a) Laut Angabe gilt

$$P(S=2|T=1) = \frac{3}{10} = \frac{P(S=2\cap T=1)}{P(T=1)},$$

durch Umformen dann $1/10$ berechnen und $8/45$ erschließen. Es gilt:

$$P(T=2\cap S=1) + P(T=2\cap S=2) + P(T=2\cap S=3) = \frac{1}{3}$$

und daher mit der Angabe

$$2P(T=2\cap S=1) = \frac{1}{3} - (T=2\cap S=2) = \frac{1}{3} - \frac{8}{45} = \frac{7}{45},$$

also

$$P(T=2\cap S=1) = P(T=2\cap S=3) = \frac{7}{90}$$

(b) Aus Tabelle

$$P(T=3|S=1) = \frac{P(T=3\cap S=1)}{P(S=1)} = \frac{1/18}{1/3} = \frac{1}{6}$$

Die Ereignisse $T=3$ und $S=1$ sind stochastisch unabhängig, wenn gilt:

$$\frac{1}{6} = P(T=3|S=1) \stackrel{!}{=} P(T=3) = \frac{1}{3}$$

Die Ereignisse sind nicht stochastisch unabhängig.

(c) Für p_X gilt:

$$p_X = 1 - \left(\underbrace{P(S=1 \cap T=1) + P(S=2 \cap T=2) + P(S=3 \cap T=3)}_{\text{"Torwart hält den Schuss"}} \right).$$

Damit ist für Werte aus Vorgabe

$$p_X = 1 - (\frac{1}{18} + \frac{1}{18} + \frac{2}{9}) = \frac{2}{3},$$

und für korrekt berechnete Werte aus TA 1.

$$p_X = 1 - (\frac{1}{5} + \frac{8}{45} + \frac{2}{9}) = \frac{2}{5}.$$

2. (a) Die Versuche müssen voneinander unabhängige Bernoulli-Experimente mit Erfolgswahrscheinlichkeit p_X sein. Alternativ: Stichprobe mit Zurücklegen.

   (b) $X_5 \sim \text{Bin}(n = 5, p = 0.4)$ – die Parameter sind also $n = 5$ und $p = 0.4$

   (c) Erwartungswert einer binomialverteilten Zufallsvariablen:

$$\mathbb{E}[X_5] = np = 5 \cdot 0.4 = 2$$

   Richtiges Ergebnis bei anderen Parametern 1.25, 3.6 oder 2.25

   (d)

$$P(X_5 = 3) = f_{\text{Bin}(n=5,p=0.4)}(3) = \binom{5}{3} 0.4^3 (1-0.4)^{5-3} = 0.2304$$

   Andere Ergebnisse: 0.0879, 0.2508, 0.2336

   (e)

$$P(X_5 \geq 4) = 1 - P(X_5 \leq 3) = 1 - F_{\text{Bin}(n=5,p=0.4)}(3) = 0.0870$$

   Richtiges Ergebnis bei anderen Parametern: 0.0156, 0.5174, 0.1657

3. (a) Ist genau die durch den **R**-Output

```
> pbinom(q=2, size=5, prob=0.68)
[1] 0.1905263
```

   gegebene Wahrscheinlichkeit.

   (b)

```
> plot(0:5,dbinom(0:5,5,0.8))
```

Aufgabe 4

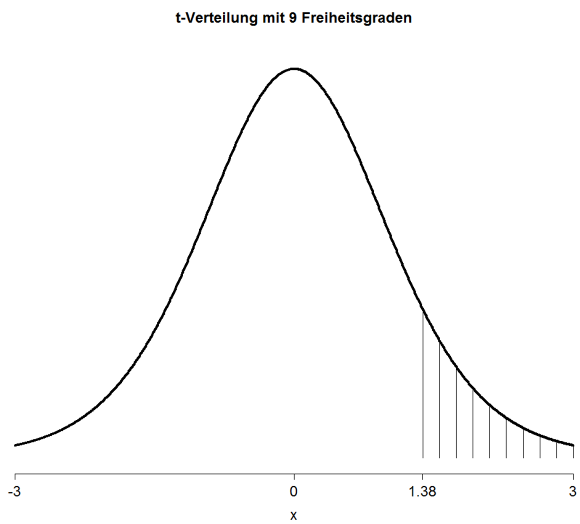
1. $s^2_{Y;8} = \frac{1}{8-1}\sum_{i=1}^{8}(y_i - \overline{y_8})^2 = \frac{1}{7}(20^2 + 4^2 + \cdots + 36^2) = \frac{1}{7}\cdot 2114 = 302$

2.

$$\begin{aligned} s_{XY} &= \frac{1}{8-1}\sum_{i=1}^{8}(x_i - \overline{x_8})(y_i - \overline{y_8}) = \frac{1}{7}\cdot 67.1 = 9.5857 \\ \rho_{XY} &= \frac{s_{XY}}{\sqrt{s^2_{X;8}s^2_{Y;8}}} = \frac{9.5857}{\sqrt{13.06\cdot 302}} = \frac{9.5857}{62.8022} = 0.1526 \end{aligned}$$

3. (a) Quantil; 5 (b) 90%; den Mittelwert

4. $K \sim \mathcal{N}(150\mu, 150^2\cdot 2) \quad \Rightarrow \quad K \sim \mathcal{N}(150\mu, 45000)$

5. 

6. Kritischer Bereich: $t_n < -t_{1-\alpha;n+m-2}$

   Kritische Schranke: $-t_{0.95;18} = -1.734$

   Teststatistik: $t_n = \frac{\overline{x}_{neu} - \overline{x}_{alt} - 0}{\sqrt{\overline{S}^2(\frac{m+n}{mn})}} = \frac{-0.5}{\sqrt{15\cdot\frac{20}{100}}} = -0.2887$

   Testentscheidung: $-0.2887 > -1.734 \quad \Rightarrow \quad H_0$ nicht ablehnen.

7.
   - größer
   - größer